

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 15:52:38 ; Search time 170.72 Seconds
(without alignments)
26.250 Million cell updates/sec

Title: US-09-331-631A-5_COPY_145_210

Perfect score: 375

Sequence: 1 KRDPQREYEDCRHCEQOE.....PORGSGRYEGEGEKQSDNP 66

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR_66:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129	34.4	509	2 S08059	alpha-globulin typ
2	121	32.3	588	1 FMCNAB	alpha-globulin B p
3	120	32.0	605	2 S06398	alpha-globulin typ
4	116.5	31.1	566	2 S22477	vicilin precursor
5	95	25.3	637	2 S35221	globulin Bgl1 prec
6	84.5	22.5	810	2 T44430	protein PVI100 [imp
7	84.5	22.5	1655	2 T13998	gene mastermind pr
8	81	21.6	646	2 D82493	conserved hypotet
9	79	21.1	1671	2 S71628	sensory transduct
10	78	20.8	604	2 T15132	ATP-dependent RNA
11	78	20.8	707	2 A48686	probable RNA helic
12	77	20.5	428	1 I36930	involutin - white
13	76	20.3	411	2 T29475	hypothetical prote
14	76	20.3	493	1 A57783	involutin - cotto
15	75.5	20.1	550	2 A46419	trophoblast-endoth
16	75	20.0	298	1 I36912	involutin S - dou
17	74.5	19.9	625	1 A34615	proliferagin - rat
18	74	19.7	236	2 T01662	globulin-1 - maize
19	74	19.7	613	2 S27770	hypothetical prote
20	73.5	19.6	669	2 JC5662	hepatoma-derived g
21	72	19.2	540	2 S21825	vicilin-like stora
22	72	19.2	573	2 A53334	globulin-1S, GLB1S
23	71.5	19.1	330	2 T25169	hypothetical prote
24	71.5	19.1	1898	1 A45973	trichomyalin - hum
25	71	18.9	544	1 I36911	involutin L - dou
26	71	18.9	777	2 B54024	protein kinase (BC
27	71	18.9	919	2 A39248	androgen receptor
28	71	18.9	930	2 T08588	hypothetical prote
29	71	18.9	1905	2 T18267	multidrug resistan

30	70	18.7	1023	2 S12519	glutactin - fruit
31	70	18.7	1053	2 T51016	related to cyc8 pr
32	70	18.7	1090	2 A41696	regulatory protein
33	69.5	18.5	255	2 A60637	mezoquite antigen
34	69.5	18.5	910	2 A34721	androgen receptor
35	69.5	18.5	911	2 B34721	androgen receptor
36	69	18.4	242	2 T29699	hypothetical prote
37	69	18.4	425	2 T18592	hypothetical prote
38	69	18.4	562	2 T26242	hypothetical prote
39	69	18.4	568	2 T26243	hypothetical prote
40	69	18.4	600	2 T18593	hypothetical prote
41	69	18.4	905	1 RGR153	regulatory protein
42	69	18.4	966	2 S25365	Cyc8 protein - yea
43	69	18.4	1038	2 T02634	rep protein homolo
44	69	18.4	1436	2 S57238	forced protein 5.4
45	69	18.4	1449	2 S57237	forced protein 5.6

ALIGNMENTS

RESULT 1

S08059

alpha-globulin type B precursor (tandem 1) - upland cotton (fragment)

N:Alternate names: seed storage protein

C:Species: Gossypium hirsutum (upland cotton)

C>Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 30-Sep-1993

C:Accession: S08059

R:Chlan, C.A.; Borroto, K.; Kamalay, J.A.; Dure III, L.

Plant Mol. Biol. 9, 533-546, 1987

A:Title: Developmental biochemistry of cottonseed embryogenesis and germination. XIX.

A:Reference number: S06398

A:Accession: S08059

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-509 <CHL>

C:Superfamily: glycinin

Query Match 34.4% Score 129; DB 2; Length 509;
Best local Similarity 35.6% Pred. No. 3.6e-06;
Matches 36; Conservative 9; Mismatches 18; Indels 38; Gaps 5;

QY 3 DPQREYEDCRHCEQOEPRLOYCQRC-----QEQQ-----RQHGRCGDLM--- 45

Db 1 DPQRR-YEECQRCRQOEERQOPQCQRCRKREYQOQOSQRFQECQCHQOEQRPEK 59

QY 46 -----NPRG-----GSGRYEGEGEKQSDNP 66

Db 60 KQCVRECKEYQENPWRGEEAEETEDEGEQCSHNP 100

RESULT 2

FMCNAB

alpha-globulin B precursor (clone C72) - upland cotton

N:Alternate names: seed storage protein; vicilin precursor

C:Species: Gossypium hirsutum (upland cotton)

C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Jul-1999

C:Accession: A30838; S06911

R:Chlan, C.A.; Pyle, J.B.; Legocki, A.B.; Dure III, L.

Plant Mol. Biol. 7, 475-489, 1986

A:Title: Developmental biochemistry of cottonseed embryogenesis and germination XVIII

A:Reference number: A30838

A:Accession: A30838

A:Molecule type: mRNA

A:Residues: 1-588 <CHL>

A:Cross-references: GB:M16891; NID:G167374; PIDN:AAA3071.1; PID:G167375

A:Experimental source: var. Coker 201

R:Chlan, C.A.; Borroto, K.; Kamalay, J.A.; Dure III, L.

Plant Mol. Biol. 9, 533-546, 1987

A:Title: Developmental biochemistry of cottonseed embryogenesis and germination. XIX.

A:Reference number: S06398

A:Accession: S06911


```
Query Match          20.8%; Score 78; DB 2; Length 604;
Best Local Similarity 28.6%; Pred. No. 0.84;
Matches 22; Conservative 11; Mismatches 26; Indels 18; Gaps 3;

OY 2 RDPQREYEDCRHCEOEPRLOYOCQ-----RRCOEQRORH-----GRGGL 44
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 5 QDPGRHRS-SDCEPRKREPRVCYNCOQPGHRSRECTERKREGRGTGGFGGACGNG 63

OY 45 MNPORGSGRYEGEER 61
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 64 GNDGFGGDCGFGGGEER 80

RESULT 11
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-707 <R00>
A:Cross-references: GB:L19948
C:Superfamily: unassigned DEAD/H box helicases; DEAD/H box helicase homology
C:Keywords: ATP; P-loop
F:329-643/Domain: DEAD/H box helicase homology <DEAD>
F:439-444/Region: nucleotide-binding motif A (P-loop)
F:443-446/Region: nucleotide-binding motif B

Query Match          20.8%; Score 78; DB 2; Length 707;
Best Local Similarity 28.6%; Pred. No. 0.97;
Matches 22; Conservative 11; Mismatches 26; Indels 18; Gaps 3;

OY 2 RDPQREYEDCRHCEOEPRLOYOCQ-----RRCOEQRORH-----GRGGL 44
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 108 QDPGRHRS-SDCEPRKREPRVCYNCOQPGHRSRECTERKREGRGTGGFGGACGNG 166

OY 45 MNPORGSGRYEGEER 61
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 167 GNDGFGGDCGFGGGEER 183

RESULT 12
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-428 <RES>
A:Cross-references: GB:M67478; NID:g176629; PIDN:AAA35405.1; PID:g176630
C:Comment: During the terminal differentiation of keratinocytes, this protein from the c
linked envelope under the plasma membrane.
C:Genetics:
A:Introns: #status absent
C:Superfamily: involucrin
C:Keywords: cornified cell envelope; duplication; epidermis; tandem repeat
F:153-386/Region: 10-residue repeats (Q-E-G-Q-[PLV]-[KE]-[LH]-[PL]-E-Q)
```

```
Best Local Similarity 29.2%; Pred. No. 0.79;
Matches 21; Conservative 17; Mismatches 18; Indels 16; Gaps 3;

OY 8 EYEDCRHCEOEPRLOY-----OCORCOEQORHGRUGDLNMP-ORGSGR 54
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 257 QDEBOLKHLQEGBGLEHLEQOEGELKHLQEGBGLEHLEQOEGO---LELPEQVYGOSK 313

OY 55 YEGBEKGSDNP 66
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 314 HLEQEEKOLEHP 325

RESULT 13
A:hypothetical protein T01D1.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C:Accession: T29475
R:Bradshaw, H.; Wohlmann, P.
submitted to the EMBL Data Library, November 1996
A:Description: The sequence of C. elegans cosmid T01D1.
A:Reference number: 220623
A:Accession: T29475
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-411 <BRA>
A:Cross-references: EMBL:U80455; PIDN:AA837887.1; GSPDB:GN00020; CESP:T01D1.6
A:Experimental source: strain Bristol N2; clone T01D1
C:Genetics:
A:Gene: CESP:T01D1.6
A:Map position: 2
A:Introns: 25/3; 304/3
C:Superfamily: gliadin

Query Match          20.3%; Score 76; DB 2; Length 411;
Best Local Similarity 29.2%; Pred. No. 0.97;
Matches 21; Conservative 12; Mismatches 21; Indels 18; Gaps 4;

OY 5 QOREYEDCRHCEOEPRLOYOCRCOE-----QOROHGRGGLMNPORG 50
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 142 QQQCQDDCAACPPQQDP--QQCCQCCCTTCGSDQYSQLIQQDTYAQPQI--PAYT 197

OY 51 GSGRYEEGEER 62
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 198 GRSEYNTAPQ 209

RESULT 14
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-493 <RES>
A:Cross-references: GB:M67477; NID:g343313; PIDN:AAA36950.1; PID:g343314
C:Comment: During the terminal differentiation of keratinocytes, this protein from th
linked envelope under the plasma membrane.
C:Genetics:
A:Introns: #status absent
C:Superfamily: involucrin
C:Keywords: cornified cell envelope; duplication; epidermis; tandem repeat
F:152-451/Region: 10-residue repeats (Q-E-G-Q-[PLV]-[KE]-[LH]-[PL]-E-Q)
```


